

University of Groningen

Opvoeden in je eentje

Bosman, Maria Hendrika

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

1993

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Bosman, M. H. (1993). *Opvoeden in je eentje: een onderzoek naar de betekenis van het moedergezin voor de onderwijskansen van kinderen*. s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

3 Onderzoeksopzet en analysestrategie

3.1 Inleiding

In dit hoofdstuk gaan we in op de onderzoeksopzet, dataverzameling en analysestrategie. Het feit dat het in dit proefschrift gaat om een reconstructie van een uit verschillende onderdelen bestaand onderzoek, brengt een gedifferentieerde beschrijving van opzet en dataverzameling met zich mee.

Vooraf merken we op dat aan de (kwantitatieve) toetsing van de probleemstellingen I en II dezelfde onderzoeksopzet en analysestrategie ten grondslag ligt, terwijl probleemstelling III om een andere, kwalitatieve benadering vraagt. Leidraad voor dit hoofdstuk vormt het onderzoeksdesign van de eerste en de tweede probleemstelling; aan de benadering van de derde probleemstelling wordt summier aandacht besteed onder verwijzing naar hoofdstuk 8. Bij de inhoudelijke beschrijving van de onderzoeksopzet en de dataverzameling moet een tweede opmerking worden gemaakt. Omwille van de overzichtelijkheid en de helderheid hebben we hiervoor steeds hier de toetsing van de *eerste* probleemstelling als voorbeeld genomen; de specifieke invulling van de variabelen en data voor de *tweede* probleemstelling vindt plaats in hoofdstuk 7, en die voor de *derde* probleemstelling in hoofdstuk 8.

Een en ander resumerend, gaan we als volgt te werk. In 3.2 en 3.3 gaan we in op de onderzoeksopzet van de beide kwantitatieve, toetsende onderdelen van het onderzoek (probleemstelling I en II). De beschrijving van variabelen in 3.4 vloeit logisch voort uit deze onderzoeksopzet; de *keuze* van variabelen heeft zoals gezegd betrekking op probleemstelling I. Een soortgelijke strategie volgen we voor de verschillende vormen van dataverzameling. In 3.5 zetten we uiteen op welke wijze in elke fase van het onderzoek de dataverzameling heeft plaatsgevonden; de beschrijving wordt ingevuld voor de dataverzameling ten behoeve van probleemstelling I. Voor een uitgebreide verantwoording van de dataverzameling ten behoeve van de probleemstellingen II en III verwijzen we wederom naar de desbetreffende hoofdstukken 7 en 8.

Nadat we in 3.6 enkele opmerkingen hebben gemaakt over de selectiviteit en representativiteit van de steekproeven, wordt in 3.7 de analysestrategie uiteengezet. Ook deze heeft betrekking op de probleemstellingen I en II. In 3.8 gaan we over op de onderzoeksbenadering van probleemstelling III en bespreken we kort onze uitgangspunten bij de kwalitatieve analyse.

In 3.9 evalueren we onze onderzoeksopzet en analysestrategie in het licht van de methodologiekritiek zoals die in 2.5 werd gepresenteerd. In 3.10 sluiten we dit hoofdstuk af.

3.2 Onderzoeksopzet

Bij de keuze van onze onderzoeksopzet hebben verschillende aspecten van de probleemstelling een rol gespeeld. In de eerste plaats gaat het in dit onderzoek om een vergelijking van schoolloopbanen en in de tweede plaats worden kinderen uit eenoudergezinnen vergeleken met kinderen uit tweeoudergezinnen.

Een onderzoek naar schoolloopbanen zal zich over een langere periode moeten uitstrekken. Dit betekent dat longitudinaal onderzoek geboden is. Peschar (1975) noemt vier typen onderzoek waarbij over langere tijd personen worden gevolgd, alsmede de relatieve voor- en nadelen van elk type. Het prospectief-achteraf onderzoek lijkt daarbij de meeste voordelen in zich te verenigen: een populatie waarover in het verleden gegevens zijn verzameld wordt nagespeurd tot in de huidige tijd. Als nadelen noemt Peschar enerzijds de afhankelijkheid van beschikbaar materiaal uit het verleden, anderzijds het feit dat personen worden gekozen op grond van een sociale situatie die in de loop van de tijd veranderd kan zijn, zodat controle hierop gewenst is.

Het eerste door Peschar genoemde nadeel was te ondervangen door de toegang tot het SMVO-bestand, een data-set van het Centraal Bureau voor de Statistiek met een longitudinaal en een cumulatief karakter (Smulders 1979). Aan het tweede genoemde bezwaar kon worden tegemoetgekomen door het opnieuw benaderen van de personen, c.q. de leerlingen uit het SMVO-bestand en hun ouders.

Een tweede aspect uit de probleemstelling dat van belang is voor de keuze van de onderzoeksopzet, is de vergelijking van gegevens van een kleine groep (kinderen uit eenoudergezinnen) met die van een veel grotere groep (kinderen uit tweeoudergezinnen) waarbij een aantal achtergrondkenmerken onder controle dienen te worden gehouden. Deze scheve verdeling stelt bijzondere eisen aan de wijze waarop deze controle kan worden gerealiseerd.

Op grond van deze overwegingen en de beschikbaarheid van de genoemde data-set werd gekozen voor een **prospectief achteraf-experiment** waarbij de experimentele groep en de controle-groep individueel op een aantal variabelen gelijkgeschakeld worden ('**matching**'). 'Matching' betekent in ons geval dat kinderen uit eenoudergezinnen individueel gekoppeld worden aan kinderen uit tweeoudergezinnen op grond van variabelen waarvan bekend is dat ze de schoolloopbaan beïnvloeden. Steeds wordt een 'paar' gevormd, bestaande uit een kind uit een eenoudergezin en een kind uit een tweeoudergezin, die verder op zoveel mogelijk kenmerken overeenstemmen, zoals sekse, leeftijd, milieu en IQ (zie model 3.1). Op deze wijze wordt de invloed van deze kenmerken geneutraliseerd, zodat de 'netto-impact' van de factor eenoudergezin op de schoolloopbaan kan worden vastgesteld. Zie voor een verdere beschrijving van de methode der gematchte paren o.a. Althausen en Rubin 1970; Peschar 1975; Clason 1977; Bosman, Louwes en Van der Meer 1980.

3.3 Beperkingen van de gekozen onderzoeksopzet

Naast voordelen kent matching ook een aantal nadelen. Individuele matching is een veeleisende methode omdat het vaak moeilijk of onmogelijk is bij een persoon met een bepaalde combinatie van kenmerken een perfecte ‘dubbelganger’ of ‘match’ te vinden. Eén van de problemen die zich bij individuele gelijkschakeling voordoen is dan ook de grote reductie van het uitgangsmateriaal, die voornamelijk optreedt bij de populatie waaruit de controle-groep afkomstig is, in het algemeen de ‘grote’ deelpopulatie.¹ Tegenover dit verlies aan breedte staat dat een dergelijke methode met betrekking tot de *dataverzameling* een hogere efficiëntiegraad biedt dan traditioneel survey-onderzoek, omdat controle op een aantal mogelijk storende variabelen *al van te voren* is ingebouwd.² Meester en De Leeuw wijzen er op dat matching, met name bij het verzamelen van follow-up informatie over bestaande cases ‘misschien wel de beste manier (is) om dat te doen’ (1983, p. 149).³

Een tweede beperking van matching is dat personen met kenmerken die sterk van het gemiddelde van een groep afwijken, een geringere kans krijgen om in het onderzoek te worden betrokken. Er kunnen immers alleen paren worden gevormd in de doorsnede van de oorspronkelijke populaties. Dit impliceert dat de gevonden verbanden kunnen zijn beïnvloed door de bijzondere samenstelling van de onderzoeksgroep en dat met generalisatie op dit punt voorzichtig moet worden omgegaan.

Desondanks zijn we van mening dat met deze semi-experimentele opzet verschillen kunnen worden aangetoond die evenzeer voor andere populaties of steekproeven gelden, maar die door de vele factoren die meespelen op een andere manier niet of moeilijk aan te tonen zijn (zie Clason 1977, p. 96). Een belangrijke factor is in dit verband het feit dat de verdelingen op relevante variabelen in beide populaties sterk van elkaar kunnen verschillen. Generalisatie heeft in feite betrekking op de vraag, of de op grond van onze resultaten getrokken conclusies ook in breder verband geldig zijn. Door op een groot aantal variabelen paarsgewijs te ‘matchen’, zijn de beide (groepen) kinderen zo gelijk als mogelijk gemaakt, gelijkter dan ze in werkelijkheid zijn.⁴ Wanneer we onder deze condities (toch) verschillen vinden, zijn deze conservatief van aard en mogen we aannemen dat de verschillen in de werkelijkheid minstens zo groot zijn. In die zin hoeven we de *geldigheid* van onze conclusies voor een grotere populatie niet in twijfel te trekken.

Waar we echter wél voor willen waarschuwen, is het zonder meer overbrengen van de **sterkte** van gevonden samenhangen naar andere steekproeven, die onder geheel andere condities zijn samengesteld. *Matching is primair een methode om verbanden te exploreren, niet om populatieparameters te schatten.*

Een laatste probleem dat we willen noemen is het verschijnsel dat bekend is geworden als ‘regressie naar het gemiddelde’. Naar aanleiding van het onderzoek van Peschar (1975) is een discussie gevoerd over de invloed van meetonbetrouwbaarheid op de match-procedure. Het zou mogelijk zijn dat een ‘paar’ kinderen die formeel gelijk gemeten kenmerken hebben in werkelijkheid van elkaar verschillen door onbetrouwbaarheid van de meetinstrumenten. Dit lijkt vooral het geval te zijn bij extreme waarden. Indien dit fenomeen zou optreden, worden een aantal paren ten onrechte gevormd en worden de resultaten daardoor vertekend (zie De Groot en Van Peet 1975; 1976; Peschar, 1976a; 1976b). Peschar toonde echter aan dat dit effect

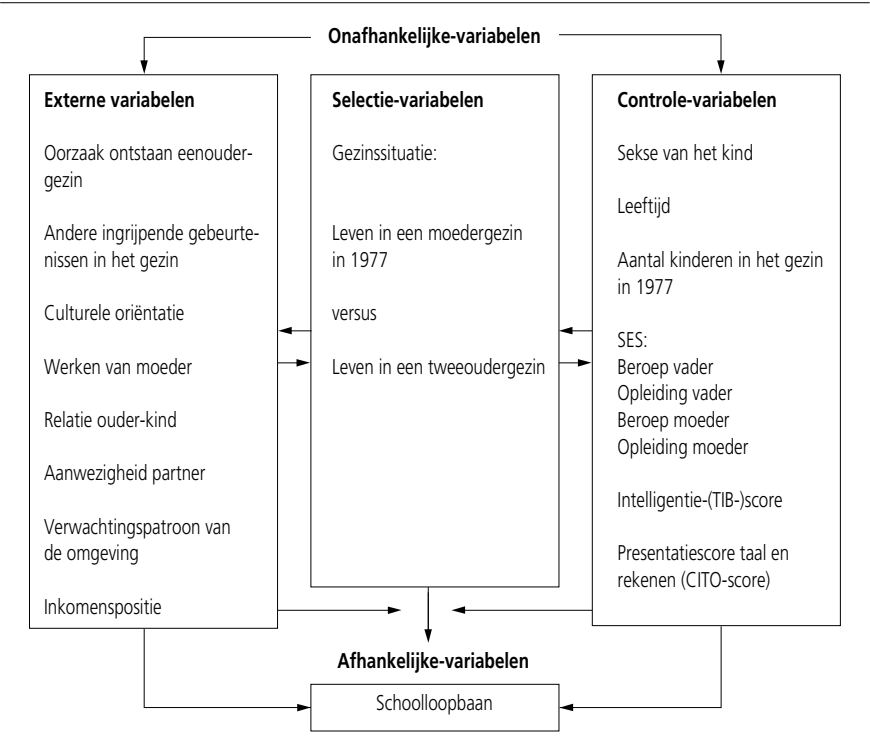
van regressie naar het gemiddelde veel geringer was dan tot dan toe gedacht was en de discussie tussen De Groot en Van Peet enerzijds en Peschar anderzijds werd beëindigd, waarbij kan worden aangetekend dat de laatste een gewonnen stelling achterlaat (Vroon 1980, p. 85).⁵

3.4 Keuze van variabelen

In Model 3.1 zijn de variabelen weergegeven die in onze analyses een rol spelen. Gegevens omtrent deze variabelen waren voor een deel beschikbaar in het SMVO-bestand en moesten voor een ander deel door ons zelf verzameld worden (zie 3.5). We geven eerst een korte toelichting op de verschillende *typen* variabelen.

In dit onderzoek hebben we ons geconcentreerd op de invloed van het leven in een eenoudergezin op de schoolloopbaan van kinderen. Met andere woorden: de afhankelijke variabele is ‘schoolloopbaan’, de onafhankelijke variabele is ‘gezins-situatie’. In werkelijkheid is er echter sprake van meerdere soorten onafhankelijke variabelen: *selectie*-, *controle*- en *externe* variabelen. (Voor de operationalisering van selectie-, controle- en afhankelijke variabelen verwijzen we naar bijlage III, tabel 3.1).

Model 3.1 Onderzoeksmodel



De *selectie-variabele* is de onafhankelijke variabele waarvan verwacht wordt dat deze de (hoofd-)oorzaak is van verschillen in de afhankelijke variabele en waarop onderscheid wordt aangebracht tussen de onderzochte kinderen. In ons geval is de selectie-variabele de gezinssituatie in 1977. Een eenoudergezin is in dit verband gedefinieerd als een gezin waarin in 1977 (het moment van de gegevensverzameling door het CBS) hetzij de moeder, hetzij de vader blijvend niet aanwezig was en waarin tenminste één thuiswonend kind in datzelfde jaar in de eerste klas van het voortgezet onderwijs zat (en dus op dat moment ongeveer 12 jaar oud was).

De *controle-variabelen* zijn de onafhankelijke variabelen waarvan bekend is dat ze verband houden met of van invloed zijn op de afhankelijke variabele en die een eventueel verband tussen het behoren tot een eenoudergezin en de schoolloopbaan zouden kunnen vertekenen. Om deze reden worden de leerlingen op dergelijke variabelen gelijkgeschakeld, zodat een eventueel verband met de selectie-variabele wordt uitgeschakeld en tegelijkertijd een mogelijk verband met de afhankelijke variabele wordt geneutraliseerd.

Op grond van literatuurstudie (Bosman en Louwes 1982) zijn de volgende kind- en gezinskenmerken als controle-variabelen gekozen: *seks* en *leeftijd* van het kind (leeftijd geïndiceerd door geboortear en geboortekwartal); een indicatie van de aanleg en/of capaciteiten van het kind, zoals gemeten met behulp van een *intelligentietest* (de zogenaamde TIB-test, een performale figuur-exclusietest) en een *prestatietoets* (CITO-toets; taal- en rekenscore opgeteld); de *gezinsgrootte* c.q. het kindertal en de *sociaal-economische status* (SES) van het gezin waartoe het kind behoort. Deze SES-score wordt geoperationaliseerd door het beroep en de opleiding van zowel de vader als de moeder van een score te voorzien (in eenoudergezinnen óók van de afwezige ouder).

De *externe variabelen* zijn onafhankelijke variabelen die wel verband (kunnen) houden met de selectie- of afhankelijke variabelen, maar waarop om verschillende redenen niet wordt gecontroleerd. Enerzijds zijn gegevens over deze variabelen soms niet beschikbaar of is de invloed ervan onbekend, anderzijds is het aantal variabelen waarop gematched kan worden om praktische redenen beperkt. Dat neemt niet weg dat de invloed van dergelijke variabelen wél in de analyses kan worden betrokken.⁶

De *afhankelijke variabele* is de variabele waarvan we willen onderzoeken of de onderzochte kinderen daarop paarsgewijs verschillen. In ons geval is dat de schoolloopbaan. Voor de operationalisatie van 'schoolloopbaan' staan ons op basis van het SMVO-bestand en de eigen gegevensverzameling een aantal indicatoren ter beschikking.

In de eerste plaats is dat het *advies van de onderwijzer voor de keuze van het voortgezet onderwijs*. In de tweede en belangrijkste plaats is dat het *onderwijs-/eindniveau* (na zeven respectievelijk negen jaar voortgezet onderwijs).⁷ Op welke wijze dit eindniveau werd bepaald, staat in Bijlage III (tabel 3.2).

Het effect van de gezinssituatie op dit eindniveau krijgt gestalte via de verschillende tussenstations in de schoolloopbaan, waarvan het 'advies voor het voortgezet onderwijs' het eerste is. Strikt genomen zouden zo vele afhankelijke variabelen te onderscheiden zijn: advies; niveau eerste klas voortgezet onderwijs; niveau tweede klas, etc. (zie bijvoorbeeld Bosman en Louwes 1982). In dit proefschrift gaat het ons echter om het *totaaleffect* van de gezinssituatie op het bereikte eindniveau en

niet om de geleidelijke opbouw daarvan. Alleen als dat van pas komt om de werking van een verklaringsmechanisme te illustreren, lichten we een enkele maal een tussenstation - voornamelijk het advies - er als afhankelijke variabele uit. Daarnaast wordt (in de toetsing van de eerste probleemstelling) nog één aanvullende karakterisering van ‘schoolloopbaan’ gebruikt, namelijk de *ononderbroken schoolloopbaan* (ontleend aan Dronkers 1987). Een leerling heeft een ononderbroken schoolloopbaan wanneer hij of zij in het vierde leerjaar gerekend vanaf 1977/78 ongeacht het niveau van het schooltype ook daadwerkelijk de vierde klas heeft bereikt, dat wil zeggen niet heeft gedoubleerd.

Tot slot van deze paragraaf geven we in schema 3.2 een voorbeeld van één van de volgens bovenstaande procedure gevormde paren. We zien dat er sprake is van ‘relatieve perfectie’; het beroep van secretaresse is ongeveer gelijk aan receptioniste, niet precies gelijk. Duidelijk blijkt ook dat met leeftijd in feite ‘klasse-leeftijd’ wordt bedoeld: het verschil tussen in december of in januari geboren zijn is dan gering, ook al is het jaartal verschillend. Een soortgelijk verschil tussen september en oktober van hetzelfde jaar zou echter door ons niet getolereerd zijn.⁸

Schema 3.2 Voorbeeld van een gematched paar

Selectie-variabelen	Moedergezin ontstaan door echtscheiding	Tweeoudergezin
Controle-variabelen		
Sekse	jongen	jongen
Leeftijd (geboortedatum)	december 1964	januari 1965
Gezinsgrootte	2 kinderen	2 kinderen
SES:		
- beroep vader	werktuigkundige	bedrijfsleider
- opleiding vader	LTS + vakopleiding	LTS + vakopleiding
- beroep moeder	receptioniste	(was) secretaresse
- opleiding moeder	ULO	ULO, steno en typen
TIB-score (range 0-33)	25	27
CITO-score (range 0-70)	52	53
Afhankelijke-variabelen		
Eindniveau na 9 jaar	MTS (mee bezig=5)	T.H. (mee bezig=8)

3.5 Dataverzameling

Zoals in 3.1 al werd uiteengezet, valt het onderzoek waarop dit proefschrift betrekking heeft uiteen in drie hoofdonderdelen, corresponderend met de probleemstellingen I, II en III. In de eerste probleemstelling is daarnaast nog een onderscheid aan te brengen tussen de toetsing van het verband tussen ‘eenoudergezin’ en ‘schoolloopbaan’ (Ia) en de toetsing van verklaringen voor dit verband (Ib). Een en ander brengt met zich mee dat er in dit onderzoek *vier verschillende datasets* zijn verzameld en/of geanalyseerd.

Het uitgangspunt voor het onderzoek werd gevormd door het SMVO-bestand,

dat gegevens van 37.000 leerlingen bevat en dat reeds in een eerder stadium werd verkend op het verband tussen 'eenoudergezin' en 'schoolprestaties' (zie Bosman en Louwes 1982). Omdat de gegevens inzake de bepaling van de precieze gezins-situatie en de sociaal-economische status van het eenoudergezin tekort schoten, werd ten behoeve van de *eerste probleemstelling* (Ia) een aanvullende schriftelijke dataverzamelingscampagne gehouden (zie 4.2.3). De onderzoeksgroep waarbinnen het verband tussen 'eenoudergezin' en 'schoolloopbaan' getoetst kon worden, bedroeg daardoor uiteindelijk 160 paren kinderen (hoofdstuk 4). Dat was de *eerste* te analyseren *dataset*.

De toetsing van probleemstelling Ib, verklaringen voor het verband tussen 'eenoudergezin' en 'schoolloopbaan', maakte het noodzakelijk dat zowel gegevens van ouders als van kinderen ter beschikking kwamen.⁹ In de schriftelijke vragenlijst ten behoeve van probleemstelling Ia was aan de ouders toestemming gevraagd voor een eventueel mondeling gesprek. Op basis van de ouders die op deze vraag bevestigend antwoordden konden opnieuw 96 paren gevormd worden: 96 moeders uit een moedergezin en 96 moeders uit een tweeoudergezin.

In verband met het 'matching-design', waarbij het (complete) paar de toetsingseenheid vormt, werd besloten pas met de interviews van de tweeoudergezinnen te beginnen nadat de gesprekken met de eenoudergezinnen gevoerd waren. De gesprekken met de kinderen waren weer afhankelijk van het feit of er een gesprek met de moeder gevoerd was, en wel in dezelfde volgorde: eerst het kind uit een moedergezin en daarna het kind uit een tweeoudergezin. Bedenken we daarbij dat onze proefpersonen over het hele land verspreid woonden en dat bovendien een groot aantal kinderen het ouderlijk huis al hadden verlaten had en elders woonden, dan mag duidelijk zijn dat er een omvangrijke, gecompliceerde en tijdrovende interviewcampagne is gevoerd.

Voor de vraaggelassen met de ouders werd een vragenlijst geconstrueerd, waarbij erop werd gelet dat enerzijds ingegaan kon worden op de specifieke situatie in eenouder- en tweeoudergezinnen, terwijl anderzijds de informatie over beide typen gezinnen zo vergelijkbaar mogelijk diende te zijn. In deze 'ouderlijst' is uitgebreid ingegaan op o.a. de volgende onderwerpen: ingrijpende gebeurtenissen in het gezin; de wijze waarop door de buitenwereld tegen het gezin wordt aangekeken; mogelijke zelfetikettering; de verstandhouding en de interactie tussen ouder(s) en kind; de financiële situatie en de onderwijs-oriëntatie binnen het gezin; en het al dan niet buitenshuis werken van de moeder. (In bijlage IV is een overzicht van de onderwerpen in deze vraaggelassen opgenomen.)

Ook bij de constructie van de vragenlijst voor de kinderen moest enerzijds bewerkstelligd worden dat de lijst geschikt was voor kinderen uit eenouder- en tweeoudergezinnen, en anderzijds dat de informatie toch goed onderling vergelijkbaar was. Deze eis van vergelijkbaarheid gold eveneens voor de informatie die door de ouders en hun kinderen werd verschaft, zodat in de 'kinderlijst' globaal gesproken dezelfde onderwerpen aan de orde kwamen als in de 'ouderlijst' en nagestreefd werd dat zoveel mogelijk dezelfde vragen werden gesteld.

Bij het hierboven genoemde aantal van 96 paren zijn dus twee keer 96 kinderen en twee keer 96 ouders betrokken, die allemaal benaderd werden voor een mondeling interview. In totaal zijn zo bijna 400 gesprekken gevoerd. Nadat de interviewcampagne was afgesloten, konden van 88 'moederparen' en 84 'kindparen' de gege-

vens worden gecodeerd en op magneetband worden overgebracht. Van 82 paren is alle informatie compleet (eenoudermoeder+tweeoudermoeder+eenouder-kind+tweeouderkind).¹⁰ De *tweede dataset* behelst daarmee in feite drie sub-bestanden.

Voor de toetsing van de *tweede probleemstelling* werd gebruikt gemaakt van een bestaand databestand, namelijk het bestand dat door Van Liere (1990) verzameld is in het kader van zijn promotie-onderzoek ‘Lastige leerlingen: een empirisch onderzoek naar sociale oorzaken van probleemgedrag op basisscholen’. Dit bestand bevat gegevens over het gedrag van kinderen en het opvoedingsgedrag van ouders die zijn verzameld bij ouders en leerlingen uit 355 tweeouder- en 51 eenoudergezinnen. ‘Matching’ van de 46 leerlingen uit moedergezinnen met leerlingen uit tweeoudergezinnen leverde 44 paren op (zie verder hoofdstuk 7). De *derde dataset* bevatte daarmee de paarsgewijze gegevens van 44 paren kinderen en 44 paren moeders.

Voor de beantwoording van de *derde probleemstelling* werd een andere weg gevolgd. Omdat de in deze probleemstelling vervatte vraag om een kwalitatieve benadering vroeg, werden 29 uitgebreide vraaggesprekken gevoerd met alleenstaande moeders met gezinsverantwoordelijkheid. Deze moeders, die benaderd werden via een zestal scholen voor basisonderwijs in de stad Groningen, hadden tenminste één kind in groep 4, 5 of 6 van de basisschool. In hoofdstuk 8 wordt uitgebreid beschreven hoe we inzake de dataverzameling te werk zijn gegaan. Daar worden ook de kenmerken van deze *vierde dataset* beschreven.

In schema 3.3 staan de verschillende onderdelen van de dataverzameling van het complete onderzoek samengevat.

Schema 3.3 Overzicht van dataverzameling

Probleemstelling	Aantal respondenten	Databron	Hoofdstuk
Ia verschillen in school-loopbanen	160 paren kinderen	Schriftelijke enquête	4
Ib verklaringen voor het verband	84 paren kinderen 88 paren moeders	Gestructureerde mondelinge vraaggesprekken	4, 5, 6,
Ila gedrag kinderen Ilb opvoedingsgedrag moeders	44 paren kinderen 44 paren moeders	Secundaire analyses dataset Van Liere	7
III Opvoedingssituatie in moedergezinnen	29 moeders	'Open' interviews	8

3.6 Selectiviteit en representativiteit

Zoals in 3.3 reeds werd opgemerkt hebben onze onderzoeksresultaten betrekking op een selectief gekozen populatie. Een beperking van ‘matching’ is immers dat personen met kenmerken die sterk afwijken van het gemiddelde van de uitgangs-

populatie, een kleinere kans hebben om in het onderzoek betrokken te worden. In de tweede plaats kunnen we door de gelijkschakeling niet spreken van twee onafhankelijke steekproeven, aangezien de keuze van kinderen uit tweeoudergezinnen volkomen afhankelijk was van de 'beschikbare' kinderen uit eenoudergezinnen.

In hoeverre de gevonden verdelingen selectief zijn wordt duidelijk uit bijlage III (tabel 3.3). Daarin worden de verdelingen van de oorspronkelijke onderzoekspopulatie op de controle-variabelen met die van de geselecteerde groep vergeleken. Deze verdelingen worden gekarakteriseerd door het gemiddelde en de standaardafwijking.

Uit de tabel blijkt dat de onderzoeksgroep op een drietal variabelen enigszins selectief is: op de TIB-(intelligentie)-score, de taalrekenscore en de opleiding van de moeder zijn de gemiddelden in de onderzoeksgroep hoger dan in de uitgangspopulatie. Aannemelijk is dat het hier in wezen om één verschil gaat, namelijk een gemiddeld hogere milieuscore van de onderzoeksgroep. Een gemiddeld hogere opleiding van de moeder brengt naar men mag aannemen immers gemiddeld hogere TIB- en taalrekenscores met zich mee (zie bijvoorbeeld Meijnen 1977).

Het gegeven dat de moeders die in ons onderzoek betrokken zijn gemiddeld hoger zijn opgeleid dan de moeders in de populatie stemt overeen met de algemene bevinding, dat in sociaal-wetenschappelijk onderzoek een grotere response te verwachten is onder hoger opgeleide personen.¹¹

3.7 Analysestrategie

Voor de toetsing van verschillen binnen de paren maken we gebruik van Wilcoxon's rangtekentoets voor gematchte paren; wanneer een variabele een dichotome verdeling heeft gebruiken we de Sign-test. De Wilcoxon-toets genereert een z-score, waarvan de interpretatie luidt: wanneer $z \geq |1.64|$ kan de nul-hypothese - geen verschil tussen beide groepen - worden verworpen ($p \leq .05$, eenzijdige toetsing).

Voor een nadere toelichting op deze werkwijze nemen we een hypothetisch voorbeeld. Daarin staat een afhankelijke variabele 'schoolsucces' centraal. De variabele heeft een range van 0 - 8, oplopend van weinig naar veel schoolsucces. In tabel 3.4 staan de paarsgewijze scores voor 'schoolsucces' voor 88 paren weergegeven. Elke cel bevat de frequentie van de combinatie van de score van een kind uit een eenoudergezin met die van het bijbehorende kind uit een tweeoudergezin. Dit betekent, dat wanneer van een paar het kind uit een eenoudergezin een *hogere* score heeft dan haar of zijn match, dit paar *boven* de diagonaal komt. *Onder* de diagonaal staan de paren waarvan het kind uit het eenoudergezin een *lagere* score behaalt dan het kind uit het vergelijkbare tweeoudergezin. *Op* de diagonaal staan de paren waarvan beide kinderen een *gelijke* score bereiken op 'schoolsucces'. Hoe verder de score van de diagonaal verwijderd is, hoe groter het verschil. Wanneer het leven in een eenouder- of tweeoudergezin geen systematische invloed zou hebben op 'schoolsucces', is het aantal paren boven en onder de diagonaal en de gemiddelde afstand tot de diagonaal gelijk.

Tabel 3.4 Verdeling van ‘schoolsucces’ van kinderen uit eenouder- en tweeoudergezinnen (hypothetisch voorbeeld)

'Schoolsucces'		Score van kind uit moedergezin								
Score van kind uit tweoudergezin		0	2	3	4	5	6	7	8	Totaal
	0		1			1				2
	2	1	3		2	3				9
	3								1	1
	4		1		5	3	1	4	2	16
	5		2		4	2	2	6	3	19
	6			1		1		4	2	8
	7					1	2	12	6	21
	8				1	2	1	3	5	12
	Totaal		1	7	1	12	13	6	29	19

We zien in dit voorbeeld dat van de 88 paren er 41 zijn waarin kinderen uit moedergezinnen een hogere score op ‘schoolsucces’ hebben dan hun ‘match’ uit een tweoudergezin; de paarsgewijze scores staan boven de diagonaal. In 27 paren bereiken de eenouderkinderen eenzelfde eindniveau als de bijbehorende kinderen uit tweoudergezinnen en in 20 paren hebben ze een lagere score. Er worden dus duidelijk meer paren *boven* de diagonaal geteld, waar het kind uit een moedergezin een hogere score heeft op ‘schoolsucces’ dan het kind uit een tweoudergezin. Met behulp van Wilcoxon’s rangtekentoets kan nu bepaald worden of deze gevonden verdeling op toeval berust; de test houdt behalve met het verschil dat per paar wordt bepaald, ook rekening met de richting en de grootte van het verschil. De verdeling blijkt in dit voorbeeld niet op toeval te berusten: de verschillen in ‘schoolsucces’ tussen kinderen uit moedergezinnen en kinderen uit tweoudergezinnen zijn significant, ten nadele van de kinderen uit tweoudergezinnen ($z = 2.72, p \leq .005$).

Om onze werkwijze bij de toetsing van *verklaringen* voor het gevonden verband tussen gezinssituatie en schoolloopbaan te verduidelijken, nemen we eveneens een fictief voorbeeld. In dit geval gaat het om de externe variabele ‘schulden’. Veronderstel dat in de literatuur een verband wordt gesuggereerd tussen het hebben van schulden en de schoolloopbaan van kinderen. Hoe zou dit verband ten nadele van kinderen uit moedergezinnen kunnen uitwerken?

De in hoofdstuk 2 besproken theoretische verklaringen impliceren doorgaans twee mogelijkheden voor de totstandkoming van een effect op de schoolloopbaan:

1. Er is sprake van een *distributed-effect*: eenoudergezinnen worden meer blootgesteld aan datgene wat in de verklaring nadelig wordt geacht voor de schoolloopbaan van kinderen. We spreken van een distributie-effect omdat de ongelijke uitkomst in schoolloopbanen in de verdeling van de randtotalen van de paren-tabel besloten ligt. Bij gelijkstelling van deze randtotalen zou het effect verdwijnen en derhalve een schijnverband blijken te zijn. In het voorbeeld van schulden zou de redenering aldus luiden: schulden zijn slecht voor de schoolloopbaan van kinderen; eenoudergezinnen hebben gezien hun inkomenspositie nu eenmaal meer schulden; ziedaar de minder gunstige schoolloopbanen van kinderen uit deze gezinnen. Bij gelijkstelling, dus wanneer eenoudergezinnen even veel of minder schulden hebben dan tweeoudergezinnen, verdwijnt het verband.

2. Er is sprake van een *interactie-effect*: eenoudergezinnen zijn *gevoeliger* voor datgene wat in de verklaring nadelig wordt geacht voor de schoolloopbaan. Het gaat mis met de schoolloopbaan van kinderen uit moedergezinnen wanneer hun gezinssituatie gepaard gaat met de in de verklaring benadrukte factor. In ons voorbeeld: het eenoudergezin krijgt negatieve effecten wanneer het in de schulden verzeild raakt. Het interactie-effect impliceert, dat bij gelijkstelling tussen eenouder- en tweeoudergezinnen het verband met de schoolloopbaan *niet* verdwijnt. ‘Evenveel’ schulden betekent in dit geval voor een eenouderkind immers iets anders dan voor een tweeouderkind.

I Het *distributie-effect* toetsen we met behulp van *gematchte differentiatie*. We gaan daarbij in twee stappen te werk. In de eerste plaats zetten we het centrale attribuut van de verklaring (bijvoorbeeld ‘schulden’) in een tabel en toetsen de paarsgewijze verdeling met behulp van Wilcoxon’s rangtekentoets (zoals in tabel 3.4).

Wanneer we geen significant verschil in de verdeling van schulden over eenouder- en tweeoudergezinnen vinden, kan van een distributie-effect geen sprake zijn en gaan we over tot de toetsing van het interactie-effect (zie II). Wanneer er wel een significant verschil in de verdeling van schulden wordt gevonden, gaan we over tot de tweede stap: het toetsen van het effect van ‘schulden’ op de schoolloopbaan.

Voor de toetsing met behulp van de gematchte differentiatie worden op grond van de parentabel drie groepen gevormd:

- groep 1 de paren waarin het eenoudergezin meer schulden heeft dan het bijbehorende tweeoudergezin (de paren boven de diagonaal van de verdeling);
- groep 2 de paren waarin beide gezinnen in het paar gelijke schulden hebben (de paren op de diagonaal);
- groep 3 de paren waarin het eenoudergezin minder schulden heeft dan het bijbehorende tweeoudergezin (onder de diagonaal).

Onder de hypothese dat ‘schulden’ het verband tussen eenoudergezin en schoolloopbaan veroorzaakt, verwachten we significante verschillen in bereikt eindniveau in groep 1, maar niet in de groepen 2 en 3. We toetsen deze verschillen in bereikt eindniveau vervolgens in de drie gevormde groepen (waarbij een groep niet kleiner mag zijn dan $n=10$).

De toetsing van het effect op de schoolloopbaan via de gematchte differentiatie kan nu drie verschillende uitkomsten geven:

1. Alleen in groep 1, dus boven de diagonaal, vinden we significante verschillen in bereikt eindniveau; op en onder de diagonaal zijn de verschillen (nagenoeg) verdwenen. De ongelijke verdeling van het kenmerk 'schulden' heeft hier een schijnverband tussen gezinssituatie en schoolloopbaan veroorzaakt; er is duidelijk sprake van louter een distributie-effect.

2. De verschillen in bereikt eindniveau ten nadele van eenouderkinderen blijven zowel boven als op en/of onder de diagonaal bestaan. Er kan dan geen sprake zijn van een distributie-effect, omdat het verband bij gelijkstelling niet is verdwenen. Wel is mogelijk, dat er sprake is van een interactie-effect; we toetsen dan nogmaals via de eenzijdige differentiatie (zie II).

3. Boven de diagonaal vinden we significante verschillen in eindniveau, op en/of onder de diagonaal zijn de verschillen weliswaar niet significant, maar nog wel substantieel. Mogelijk is er sprake van een gecombineerd distributie- en interactie-effect. We toetsen daarom dit laatste nogmaals via de eenzijdige differentiatie (zie II).

II Het *interactie-effect* toetsen we via *eenzijdige differentiatie*. Dat wil zeggen dat we nagaan of de verschillen in schoolloopbaan hoofdzakelijk optreden binnen die paren waarin het eenoudergezin het in de verklaring benadrukte kenmerk in hoge mate bezit. In ons voorbeeld betekent dit, dat we de groep eenoudergezinnen opsplitsen in een groep met veel en een met weinig schulden. Wanneer van een interactie-effect sprake is, zal blijken dat de verschillen in eindniveau zich concentreren binnen de groep kinderen uit moedergezinnen met veel schulden.

We zien dat beide strategieën hun sterke en hun zwakke kanten hebben. De eenzijdige differentiatie kan een door de verklaring voorspeld effect aantonen, maar is op zichzelf niet in staat om te onderscheiden tussen distributie- en interactie-effecten. Dat wil zeggen: deze strategie staat niet toe te melden dat het effect onder gelijke omstandigheden zou verdwijnen.

De strategie van de gematchte differentiatie daarentegen is wel geschikt om het distributie-effect en een eventueel schijnverband bloot te leggen, maar is niet het meest geschikt om een interactie-effect zichtbaar te maken. We kunnen dit illustreren aan de hand van ons voorbeeld van de schulden. De strategie van de eenzijdige differentiatie verwijst naar de *absolute* hoogte van de schulden, terwijl het bij de gematchte differentiatie in feite gaat om de *relatieve* hoogte van de schulden, d.w.z. 'hoog' in vergelijking met wat in een overeenkomstig tweeoudergezin wordt aangetroffen. Het is duidelijk dat beide strategieën in combinatie het grootste inzicht verschaften.¹²

In schema 3.4 zijn de verschillende mogelijkheden nog eens samengevat. Hieruit blijkt dat ook een gecombineerd distributie- en interactie-effect tot de mogelijkheden behoort. Met 'oorzakelijk kenmerk' wordt het in de verklaring benadrukte kenmerk bedoeld; in het door ons gebruikte voorbeeld was dat dus 'schulden'.

Voor wat betreft de keuze tussen beide strategieën zijn er drie redenen om in voorkomende gevallen voor slechts één van beide te kiezen:

- 1 wanneer de randtotalen niet verschillen, wordt alleen via de eenzijdige differentiatie getoetst;

- 2 wanneer bij gematchte differentiatie het verband op de diagonaal is verdwenen (of in de andere richting ligt), kan er ook geen sprake zijn van een interactie-effect;
- 3 wanneer de theorie zelf duidelijk aanwijst of het om een distributie- of een interactie-effect gaat.

Schema 3.4 Overzicht analysestrategieën

Geen effect		DISTRIBUTIE-EFFECT		Schijnverband
		NIET	WEL	
NIET	INTERACTIE EFFECT	Toetsing via gematchte differentiatie toont geen een-duidig verband met het oorzakelijk kenmerk èn	Toetsing via gematchte differentiatie toont een een-duidig verband met het oorzakelijk kenmerk èn	
		Toetsing via eenzijdige differentiatie toont evenmin een duidelijke localisering van het effect in de eenouder-gezinnen met het oorzakelijk kenmerk	Het effect verdwijnt voor de paren die het kenmerk in gelijke mate bezitten	
WEL		Toetsing via gematchte differentiatie toont geen een-duidig verband met het oorzakelijk kenmerk maar	Toetsing via gematchte differentiatie toont een een-duidig verband met het oorzakelijk kenmerk maar	
		Toetsing via eenzijdige differentiatie toont wèl een duidelijke localisering van het effect in de eenouder-gezinnen met het oorzakelijk kenmerk	Het effect is niet verdwenen voor de paren die het kenmerk in gelijke mate bezitten èn Toetsing via eenzijdige differentiatie toont wèl een duidelijke localisering van het effect in de eenouder-gezinnen met het oorzakelijk kenmerk	
Interactie-effect		Gecombineerd effect		

3.8 De kwalitatieve benadering

Het onderzoeksdesign van de twee eerste delen van dit onderzoek, waarin de eerste en de tweede probleemstelling worden getoetst, is gebaseerd op een kwantitatieve vergelijking tussen kinderen en moeders uit eenouder- en tweeoudergezinnen. De verbanden tussen ‘eenoudergezin’ en ‘schoolloopbaan’ en tussen ‘eenoudergezin’ en ‘opvoedingsgedrag’ worden getoetst met behulp van grootschalig, statistisch onderzoek. Onderzoek naar *de mechanismen via welke deze verbanden tot stand komen* vraagt evenwel om een andere benadering. Bovendien is het ook voor een goede interpretatie van het uit de cijfermatige toetsing gerezen beeld noodzakelijk

dit beeld aan te vullen met een meer kwalitatieve analyse van de situatie in moeder-gezinnen. Dit vraagt om een meer kleinschalig onderzoek, waarin ruimte bestaat voor heterogeniteit en nuances. Dit type onderzoek wordt ook wel aangeduid met de term *'intensief'* onderzoek (Van der Veen, 1988; Te Grotenhuis, 1993).

Van der Veen (1988) hanteert het onderscheid tussen intensief en extensief onderzoek in navolging van Harré en De Waele (1976; zie ook Schuyt 1986) als alternatief voor het meer gehanteerde onderscheid tussen kwalitatief en kwantitatief onderzoek. Kenmerkende aspecten van intensief onderzoek zijn volgens Van der Veen:

- en beperkt aantal cases wordt onderzocht op veel verschillende en - dat is de vooronderstelling - met elkaar samenhangende eigenschappen;
- het onderzoek levert geen waarschijnlijkheidsuitspraken op, maar beschrijvende en verklarende uitspraken;
- generalisatie vindt plaats door middel van abstractie;
- in intensief onderzoek wordt gegeneraliseerd naar het onderzochte verschijnsel in plaats van naar de populatie zoals dat in extensief onderzoek het geval is.

Belangrijke pleiters voor de intensieve benadering in sociaal-wetenschappelijk onderzoek zijn Glaser en Strauss, die in hun *'Discovery of Grounded Theory; strategies for qualitative research'* (1971) baanbrekend werk verrichtten op het terrein van het kwalitatieve onderzoek. In tegenstelling tot de door hen gehanteerde methode speelt in intensief onderzoek theorie wel een belangrijke rol bij de voorbereiding van het onderzoek.

'Juist door middel van een goede theoretische voorbereiding kan beter worden waargenomen: theorie is de bril die een onderzoeker nodig heeft om de sociale werkelijkheid te kunnen zien' (Van der Veen 1988, p. 144).

Ondanks dit belangrijke verschil tussen kwalitatief en intensief onderzoek, biedt het werk van Glaser en Strauss goede aanknopingspunten voor de verzameling, ordening en analyse van meer kwalitatieve gegevens. De methodologie van het kwalitatieve onderzoek is namelijk veel minder uitgewerkt dan die van de kwantitatieve benadering, maar is tegelijkertijd sterk in ontwikkeling (zie bijvoorbeeld Wester 1987; Strauss en Corbin 1990). Een goede ingang vormt ook het recente werk van Gilgun, Daly en Handel (1992), waarin betoogd wordt dat juist gezinsonderzoek om een kwalitatieve benadering vraagt. (Een uitmuntend voorbeeld van kwalitatief gezinsonderzoek op een met het onze verwant onderzoeksterrein is de studie van Clark: *'Family Life and School Achievement: why poor black children succeed or fail'* (1983).)

Voor de methodologie van het *kwalitatieve interview* biedt McCracken (1988) nuttige informatie. Voor de analyse van de gegevens kan worden aangesloten bij het principe van *'vergelijking'* zoals bijvoorbeeld Clark (1983) dat hanteert, en dat grote overeenkomst vertoont met de *'constant comparison'* van Glaser en Strauss (1971). De categorieën waarop de gegevens vergeleken worden, zijn gebaseerd op het conceptuele schema dat ten grondslag ligt aan de niet-gestructureerde vragenlijst (zie hoofdstuk 8).

Voor een uitgebreide verantwoording van de keuze en de wijze van benadering van de respondenten, de steekproef en de dataverzameling verwijzen we naar hoofdstuk 8. Daarop vooruitlopend willen we benadrukken dat door in dit onder-

zoek een dergelijke gedetailleerde waarneming van de sociale werkelijkheid op te nemen, de mogelijkheden tot interpretatie en betekenisgeving van de resultaten enorm worden vergroot. Juist in *de combinatie* van de kwalitatieve en de kwantitatieve benadering schuilt ons inziens een belangrijke innovatieve stap in dit type sociologisch onderzoek. (Zie voor een recent voorbeeld van deze combinatie Te Grotenhuis 1993).¹³

3.9 Confrontatie met de methodologiekritiek

In dit hoofdstuk is de analysestrategie uiteengezet met behulp waarvan we de onderscheiden probleemstellingen willen toetsen. Wanneer we ons onderzoeksdesign en onze werkwijze terugkoppelen naar de methodologische aanbevelingen in 2.5, dan kunnen we de volgende conclusies trekken.

Ten aanzien van de *steekproeftrekking* wordt aan de aanbeveling zoals die door Blechman (1982) wordt verwoord volledig recht gedaan: het SMVO-bestand, maar ook de dataset van Van Liere zoals die in hoofdstuk 7 wordt geanalyseerd, zijn representatief voor de ‘general family population in a setting typically frequented by the general family population’ (p. 189). In beide gevallen gaat het om representatieve steekproeven uit een populatie van respectievelijk leerlingen van het voortgezet en het basisonderwijs, waarbij vooraf geen selectie op gezinstypen heeft plaatsgevonden.

Met het oog op de *onafhankelijke variabele* ‘gezinssituatie’, aanvankelijk onnauwkeurig en niet-valide gemeten in het SMVO-bestand, is in een omvangrijke eigen data-campagne aanvullend materiaal verzameld ten aanzien van de sekse van het gezinshoofd, de oorzaak en het tijdstip van ontstaan van de gezinssituatie, en de aanwezigheid van een eventuele partner. Deze data-verzameling is ook benut voor een nauwkeurige vaststelling van de SES-indicatoren van de beide ouders, dus ook van de afwezige partner en de eventuele nieuwe partner.

Voor de *controle op relevante kind- en gezinskenmerken*, met name voor de zo belangrijke sociaal-economische status van het gezin, zijn daarmee de mogelijkheden sterk vergroot. Doordat deze controle plaatsvindt met behulp van *individuele matching*, wordt bovendien zowel het eventuele verband van deze kenmerken met de onafhankelijke variabele geneutraliseerd als ook hun verband met de afhankelijke variabele onder controle gehouden. Van belang is tevens de beschikbaarheid van *objectieve intelligentie- en prestatiescores*, die ons in staat stellen de invloed van individuele begaafdheid op de schoolloopbaan onder controle te houden. Door hierop te controleren, hebben de kinderen als het ware een ‘gelijke start’ bij het begin van de schoolloopbaan.¹⁴

De *afhankelijke variabele* zelf is met behulp van een jaarlijks herhaalde meting objectief vastgesteld, hetgeen dit gegeven een grote betrouwbaarheid verleent. Dit is de vrucht van het longitudinale karakter van het SMVO-bestand, dat in dit opzicht voorziet in een belangrijke aan de *onderzoeksopzet* te stellen eis. We moeten hier echter aan toevoegen dat dit longitudinale karakter slechts ten dele van toepassing is op de gezinsgeschiedenis. De informatie daaromtrent is retrospectief verzameld.

Een belangrijke uitbreiding van de methodologische mogelijkheden schuilt in

de gehanteerde analysestrategie. In de onderzoeksliteratuur wordt dikwijls gewezen op het gevaar van schijnverbanden. Juist de mogelijkheid om *distributie-effecten* te kunnen onderscheiden van *interactie-effecten* verschaft ons een middel om de aard van de effecten op hun juiste waarde te kunnen schatten.

Een tweede in onze ogen belangwekkend methodologisch winstpunt betreft de combinatie van *kwalitatieve* analyses en *kwantitatieve* toetsing, zoals die in dit onderzoek wordt gehanteerd. Het is op zichzelf reeds verbazingwekkend dat juist op dit onderzoeksterrein zo weinig kwalitatieve studies zijn uitgevoerd (zie ook Amato 1993b, p. 53). De combinatie van een kwantitatieve en een kwalitatieve benadering is evenwel nog veel zeldzamer. Toch is het juist deze combinatie, die een goede interpretatie van het uit de cijfermatige toetsing gerezen beeld mogelijk maakt.

Alhoewel ook onze werkwijze beperkingen kent, zoals al in 3.3 werd opgemerkt, zijn we alles overziend van mening dat we in onze onderzoeksofzet en analysestrategie in belangrijke mate tegemoetkomen aan de voornaamste methodologische eisen en dat we in dit opzicht wellicht een enkele stap vooruit kunnen maken. Tot besluit van dit hoofdstuk schetsen we hoe en waar de geschetste aanpak in de beantwoording van de probleemstellingen tot uitdrukking zal komen.

3.10 Beantwoording van de onderzoeksvragen en toetsing van de hypothesen

In de nu volgende hoofdstukken worden de hypothesen getoetst, zoals die worden afgeleid uit de verschillende aspecten van de probleemstelling. Daarbij wordt als volgt te werk gegaan.

In hoofdstuk 4 gaat het om de eerste (Ia) en een deel van de tweede probleemstelling (Ib): de toetsing van verschillen in schoolloopbanen tussen kinderen uit eenouder- en tweeoudergezinnen en de invloed van een tweetal factoren op het verband tussen 'gezinssituatie' en 'schoolloopbaan'. Ook wordt in hoofdstuk 4, in een eerste verkenning van verklarende factoren, aandacht besteed aan enkele meer kwalitatieve gegevens, uit de vraaggesprekken met moeders en kinderen.

In de hoofdstukken 5 en 6 gaat het om probleemstelling Ib en wordt een tweetal verklaringen getoetst voor het verband tussen het behoren tot een eenoudergezin en de schoolloopbaan. Beide hoofdstukken zijn als volgt opgebouwd. In de eerste paragraaf worden de theoretische en empirische noties weergegeven die de desbetreffende verklaring onderbouwen of ondersteunen. In de tweede paragraaf vindt de operationalisatie van de variabelen plaats en worden de hypothesen geformuleerd. In deze hypothesen wordt het verband tussen gezinssituatie en schoolloopbaan als uitgangspunt genomen, dat wil zeggen dat de hypothetische invloed van een 'derde' factor op dit verband wordt gespecificeerd. Als we nog even teruggrijpen op het voorbeeld van de schulden, dan zou een hypothese als volgt kunnen luiden:

De verschillen in schoolloopbaan tussen kinderen uit moedergezinnen en kinderen uit tweeoudergezinnen zijn het grootst voor kinderen uit moedergezinnen met meer schulden dan de vergelijkbare tweeoudergezinnen.

Deze formulering heeft betrekking op toetsing van het distributie-effect; voor het interactie-effect zou het gaan om *kinderen uit moedergezinnen met in absolute zin*

veel schulden. Toetsing van deze hypothesen vindt plaats in de derde paragraaf. Zowel de gematchte als de eenzijdige differentiatie kunnen daarbij worden toegepast volgens de hiervoor genoemde criteria. In de vierde paragraaf tenslotte volgt een korte conclusie.

In hoofdstuk 7 worden de hypothesen afgeleid uit de probleemstellingen IIa en IIb getoetst. Dit gebeurt volgens dezelfde analyse-strategie als in de hoofdstukken 5 en 6. Omdat het om secundaire analyses gaat, wordt in dit hoofdstuk ook de noodzakelijke aandacht besteed aan de oorspronkelijke dataverzameling en aan het karakter van het databestand.

Hoofdstuk 8 is gewijd aan probleemstelling III en heeft een kwalitatief karakter. Omdat de gegevens voor deze kwalitatieve analyse speciaal voor dit onderzoek verzameld zijn, wordt ook in dit hoofdstuk om te beginnen aandacht geschonken aan de dataverzameling en -analyse. In dit hoofdstuk is voor het eerst geen sprake van een vergelijking tussen eenouder- en tweeoudergezinnen, maar van een nadere verkenning van de opvoedingssituatie in alleen moedergezinnen. Centraal staat daarbij de vraag naar het optreden van mogelijke opvoedingsspanningen.

1. De reductie van het materiaal is afhankelijk van het aantal controle-variabelen en de gewenste perfectie van de matching. Hoe meer kenmerken men in de match-procedure betreft, hoe groter de uitval; anderzijds zijn er meer mogelijkheden om paren te vinden wanneer we grotere afwijkingen op de controle-variabelen accepteren.

Met betrekking tot de kwestie of de onderzoeksresultaten beïnvloed zullen worden door het opnemen van niet-perfecte 'matches' in het onderzoek, bestaan verschillende opvattingen (zie Gadourek 1972). Peschar toonde echter aan: *'There is no evidence for the general existence of a 'Freedman-effect', relating the completeness of matching with disappearing significance'* (1977, p. 227).

Zelf hebben wij - mede op grond van de uitspraak van Peschar - gekozen voor het compromis van de 'haalbaarheid' van een zo groot mogelijke perfectie. Dat impliceert dat wij sterk aan perfectie hebben vastgehouden, maar in gevallen waarin hetzij de theorie hetzij het gezond verstand dat toeliet, ruimere grenzen hebben getolereerd. Een voorbeeld van het laatste is het verschil in leeftijd tussen beide kinderen in schema 3.2. Het getolereerde verschil doet de 'gelijkheid' van beide kinderen geen geweld aan, terwijl wanneer strak aan 'gelijk geboortejaar' was vastgehouden geen match was gevonden. Wellicht is het daarom beter van 'relatieve gelijkheid' te spreken.

2. McNemar (1969) beschrijft enkele voorwaarden en voordelen van de methode der gematchte paren. *'The use of paired individuals is considered good experimental technique - it cannot be said that a found difference between means for the variable being studied may be due to a lack of comparability of the two groups with respect to the matching variables. The use of paired individuals has a statistical as well as an experimental advantage in that the sampling error of the difference between means is (thereby) reduced without the necessity of increasing the number of cases. If pairing produces an r of .75, the reduction in Sdm is equivalent to that achieved by quadrupling the number of cases when the random method of forming groups is employed'* (p. 95).

McNemar besluit met te beschrijven hoe het gebruik van 50 paren dezelfde informatie zou hebben opgeleverd als een survey-onderzoek met een experimentele groep van 1000 en een controle-groep van 1000 proefpersonen!

3. Meester en De Leeuw spreken voor wat betreft een beschrijving op 'paar-'niveau van 'een quasi-experimentele versie van de beroemde gescheiden opgevoede één-eiige tweelingen' (1983, p.149).
4. Uiteraard is het ook mogelijk interveniërende variabelen constant te houden door op covariaten te controleren. Door de geringe overlap van de verdelingen op relevante variabelen vraagt dit echter om zeer omvangrijke steekproeven van eenoudergezinnen. Dergelijke steekproeven zijn in Nederland niet beschikbaar.
5. In ons geval zou regressie naar het gemiddelde kunnen optreden ten aanzien van de scores op de TIB-test (ruimtelijke intelligentie) en de prestatietoetsen van het CITO. Nadere analyse wees uit dat de voor regressie-effecten gecorrigeerde scores zeer geringe verschillen met de oorspronkelijke scores vertonen.

De betrouwbaarheid van de TIB-test (figuur-exclusietest, zie Van der Flier, 1980) wordt geschat op .79. Voor het berekenen van de invloed van het regressie-effect maken we gebruik van de formule die door De Groot en Van Peet (1975) en Peschar (1976) gehanteerd werd. De range van de TIB-test loopt van 0-33 TIB-punten.

$X_t(\text{weeouder})=22.7$

$X_e(\text{enouder}) = 21.8$

Onder de veronderstelling dat $r_{xx} = .21$ is de invloed van het regressie-effect:

$(1-r_{xx})(X_t - X_e) = (.79)(.9) = 0.71$ punten TIB-score.

De betrouwbaarheid van de CITO-toetsen mag op .80 worden gesteld.

De range van de opgetelde taal-rekenscore is 0-70.

$X_t(\text{weeouder}) = 42.9$

$X_e(\text{enouder}) = 38.1$

Met behulp van dezelfde formule levert dit een regressie-effect op van 3.8 punten taal-rekenscore. Het lijkt ons onwaarschijnlijk dat een dergelijk klein verschil onze onderzoeksresultaten zal beïnvloeden.

6. Bij de externe variabelen in het schema ontbreekt de 'duur sinds het ontstaan van het eenoudergezin'. Omdat de kinderen op 12-jarige leeftijd gematched zijn, zijn eventuele effecten van 'duur' geneutraliseerd, als het ware glad gestreken door de matching. Het effect van 'duur' is in onze onderzoeksopzet dan ook niet te toetsen. Overigens moet bedacht worden dat achter de 'duur sinds het ontstaan van het eenoudergezin' ook de leeftijd van het kind bij ontstaan schuilgaat. Het ontrafelen van de effecten van deze beide factoren en eventuele interactie-effecten vraagt om aanzienlijke steekproefkaders.
7. De beschikbare versie van het SMVO-bestand bevatte gegevens over zes schooljaren sinds september 1977. Eigen aanvullende schriftelijke en mondelinge dataverzameling (zie 3.5) stelde ons in staat, een 'eindniveau' na zeven jaar respectievelijk negen jaar voortgezet onderwijs te berekenen. Van 'eindniveau' wordt hier gesproken, omdat we met behulp van deze gegevens beter in staat zijn voor de leerlingen die het onderwijs verlaten hebben een adequate score te berekenen, namelijk met inachtneming van het behalen van een diploma. Op welke wijze dit eindniveau werd bepaald staat weergegeven in bijlage III (tabel 3.2).
8. Bij de matching op leeftijd in maanden werd een verschil van enkele maanden getolereerd (bijvoorbeeld februari-april). Hierbij is gelet op de toelatingsleeftijd voor het lager onderwijs: op 1 oktober 6 jaar zijn. Dit betekent dat de volgende twee meisjes niet aan elkaar gematched werden: het eerste geboren op 30 september 1964, het tweede op 10 oktober 1964. Dit omdat het eerste meisje op de lagere school heeft gedoubleerd en het andere niet.
9. Alhoewel wij in geval van tweeoudergezinnen en eenoudergezinnen ontstaan door echtscheiding graag met beide ouders hadden gesproken, hebben we ons noodgedwongen beperkt tot de moeders.
10. Wellicht ten overvloede willen we opmerken dat de 'kinderen' in ons onderzoek - in 1977 ongeveer 12 jaar - ten tijde van de interviews de leeftijd van 21 jaar ruimschoots bereikt hadden. Net als iedereen blijven ook zij echter kinderen van hun ouders, en het is in die betekenis dat wij het woord 'kinderen' gebruiken.
11. Daarbij kan opgemerkt worden dat er binnen onze onderzoeksopzet verschillende malen sprake is geweest van een mogelijkheid tot weigering en non-response. In de eerste plaats kon men reeds bij de gegevensverzameling door het CBS in 1977 aangeven niet verder benaderd te willen worden. In de tweede plaats kon men de schriftelijke vragenlijst niet invullen en/of terugsturen. Ten derde kon de vraag in de vragenlijst, waarin de medewerking van respondenten werd gevraagd voor een mondeling gesprek, negatief beantwoord worden. Tenslotte bood het aanschrijven van respondenten voor het daadwerkelijk voeren van een dergelijk gesprek de mogelijkheid dit (alsnog) te weigeren.
12. Hierbij dient te worden aangetekend dat het *gecombineerde effect* idealiter aangetoond zou moeten worden binnen de paren die een ongunstige score op de oorzakelijke variabele gemeenschappelijk hebben, d.w.z. in één specifieke cel op de diagonaal. In de praktijk

bevat deze cel echter zulke kleine aantallen, dat het significantie-criterium hier geen uitsluitel biedt. Daarom kiezen we voor een aanvullende eenzijdige differentiatie om het gecombineerde effect aan te tonen.

13. In een onderzoek naar uitvalmotieven van deelnemers aan de beroepsgerichte volwasseneneducatie werd door ons al eerder op bescheiden wijze gebruik gemaakt van de combinatie van een kwantitatieve en een kwalitatieve benadering (Bosman 1990, 1992).
14. Uit deze formulering valt af te leiden dat eventuele nadelige effecten van 'gezinssituatie' die vóór de leeftijd van 12 jaar kunnen zijn opgetreden, door de 'matching' zijn geneutraliseerd.